Geophys. J. Int. (2003) **153,** 389–402

Discrete tomography and joint inversion for loosely connected or unconnected physical properties: application to crosshole seismic and georadar data sets

M. Musil, H. R. Maurer and A. G. Green

Institute of Geophysics, ETH Zürich, Switzerland. E-mail: green@aug.ethz.ch

Accepted 2002 October 14. Received 2002 October 14; in original form 2002 March 20

SUMMARY

Tomographic inversions of geophysical data generally include an underdetermined component. To compensate for this shortcoming, assumptions or *a priori* knowledge need to be incorporated in the inversion process. A possible option for a broad class of problems is to restrict the range of values within which the unknown model parameters must lie. Typical examples of such problems include cavity detection or the delineation of isolated ore bodies in the subsurface. In cavity detection, the physical properties of the cavity can be narrowed down to those of air and/or water, and the physical properties of the host rock either are known to within a narrow band of values or can be established from simple experiments. Discrete tomography techniques allow such information to be included as constraints on the inversions. We have developed a discrete tomography method that is based on mixed-integer linear programming. An important feature of our method is the ability to invert jointly different types of data, for which the key physical properties are only loosely connected or unconnected. Joint inversions reduce the ambiguity in tomographic studies. The performance of our new algorithm is demonstrated on several synthetic data sets. In particular, we show how the complementary nature of seismic and georadar data can be exploited to locate air- or water-filled cavities.

Key words: crosshole tomography, georadar, joint inversion, seismic, traveltime.

1 INTRODUCTION

Tomography is a widely used geophysical technique for determining 3-D spatial variations of physical properties (e.g. Menke 1989). Energy generated by numerous seismic or electromagnetic sources propagates through the media of interest to be registered by arrays of sensors placed at the surface or within boreholes. Tomographic inversion of the recorded data provides the required subsurface information. Successful applications of tomography have been reported for whole-earth investigations, hydrocarbon and mineral exploration, engineering projects and natural hazard studies.

Because source and receiver arrays are usually restricted to the surface or a small number of shallow boreholes, critical parts of the target media may be only sparsely sampled, resulting in ambiguities in the tomographic inversions. To compensate for limitations of the recorded data, additional constraints are generally required. One option is to assume that spatial variations of the subsurface physical properties are smooth. This may be implemented using an inversion algorithm that minimizes the curvature of the model space (Constable *et al.* 1987). A potential disadvantage of such a procedure is that the resultant images may be blurred and important small-scale features may remain unresolved. Another way to compensate for sparse data is to introduce *a priori* information in the form of damping. In this approach, model parameters are not allowed to deviate greatly from a given starting model (Marquardt 1970). Clearly, this requires that the starting model should be a close representation of the true subsurface structure.

Although smoothing and damping are powerful mathematical tools, it is much better to minimize the ambiguities by applying appropriate data constraints. This has led to the concept of joint inversions, whereby different types of data are inverted simultaneously (Vozoff & Jupp 1975). A necessary requirement for a joint inversion is to have a factor that is common to the two data sets. The most straightforward approach is to invert data sets that are sensitive to the same physical property. For example, direct-current electrical resistivity and electromagnetic data are both sensitive to electrical resistivity. A variety of studies have demonstrated the substantial reduction in ambiguity that may result from joint inversions (Vozoff & Jupp 1975; Jupp & Vozoff 1977; Raiche *et al.* 1985; Sandberg 1993; Maier *et al.* 1995; Schmutz *et al.* 2000).

Jointly inverting data sets that are sensitive to different physical properties is a more difficult problem. Coupling of the two data sets must involve common structural elements. In 1-D applications, the common elements may be layer thicknesses (e.g. Hering *et al.* 1995). This concept can be extended to 2- and 3-D data sets, as long as the targets can be represented by different physical models

with common geometries (Lines *et al.* 1988). Haber & Oldenburg (1997) employ a Laplacian operator in combination with a non-linear structure operator to relate the curvatures of models derived from coincident seismic and geoelectric data sets (see also Zhang & Morgan 1997). By minimizing data misfits and differences between the seismic velocity and electrical resistivity structures, they achieve a joint inversion.

Besides smoothing, damping and joint inversion, there exists a further option for reducing model ambiguity: *a priori* knowledge may enable the model parameters to be restricted to a few narrow ranges of values. An important and highly topical example would be cavity (e.g. caves, mines and tunnels) detection, in which the physical properties of the cavity are known (either those of air or water) and those of the host material can be assumed to lie within a well-defined restricted interval. If this type of information can be included in an inversion algorithm, the model space and thus the ambiguities can be significantly reduced relative to standard least-squares inversions that allow the model space to be continuous and unlimited.

Discrete tomography is a possible option for tackling problems characterized by variables that can only assume values within very limited ranges (Herman & Kuba 1999). This tomographic method has been used to map molecules in discrete lattices, reconstruct the shapes and dimensions of industrial parts (Browne *et al.* 1998) and determine approximate binary images from discrete X-rays (Gritzmann *et al.* 2000). To our knowledge, discrete tomography has not been applied to seismic and georadar traveltime tomography.

In this contribution, we present a new discrete tomography algorithm based on mixed-integer linear programming (MILP). An important advantage of our MILP formulation is that it lends itself naturally to the concept of joint inversion. Indeed, it allows all options for reducing ambiguities (i.e. smoothing, damping, joint inversion and discrete parameter intervals) to be considered simultaneously.

We begin by reviewing traveltime tomography and the commonly employed least-squares L_2-norm minimization procedure (the conventional approach). Since our MILP algorithm is based on linear programming and L_1-norm minimization, the necessary theoretical background for these concepts are outlined. After showing how the MILP technique can be applied to discrete tomography problems, we present an extension that makes it amenable to joint inversion problems. The possibilities and limitations of our approach are demonstrated on synthetic traveltime data generated from simple models with relatively high velocity contrasts. In a second suite of examples, we simulate realistic full-waveform seismograms and radargrams for typical cavity detection problems. In these latter examples, we deal with very high-velocity contrasts that generally cause difficulties in conventional tomographic inversions.

2 GENERAL TRAVELTIME TOMOGRAPHY

Traveltimes of first-arriving seismic or georadar waves can be used to derive velocity models of the subsurface (e.g. Nolet 1987, and references therein). Since first breaks are readily identifiable and their arrival times are easy to pick from high-quality data, this technique has been used successfully in surface seismic refraction (e.g. Zelt & Smith 1992; Lanz *et al.* 1998), seismic crosshole (e.g. Dyer & Worthington 1988; Chapman & Pratt 1992; Williamson *et al.* 1993; Maurer & Green 1997) and georadar (Musil *et al.* 2002) crosshole investigations.

The traveltime t of a seismic or georadar wave travelling along a ray path S through a 2-D isotropic medium can be written as

$$t = \int_S u(\mathbf{r}(\mathbf{x}, \mathbf{z}))\, dr, \tag{1}$$

where $u(\mathbf{r})$ is the slowness (the reciprocal of velocity) field and $\mathbf{r}(\mathbf{x}, \mathbf{z})$ is the position vector. The slowness field $u(\mathbf{r})$ is represented by M cells, each having a constant slowness u_j ($j = 1, \ldots, M$), so the ith traveltime can be written as

$$t_i = \sum_{j=1}^{M} l_{ij} u_j = \mathbf{L}_i \mathbf{u}, \tag{2}$$

where l_{ij} denotes the portion of the ith ray path in the jth cell. To determine the matrix $\mathbf{L}$, calculation of ray paths in 2-D media is required. In strongly heterogeneous media, this can be achieved by first computing the traveltime fields using a finite-difference approximation of the eikonal equation and subsequently reconstructing the ray paths (e.g. Lanz *et al.* 1998).

Eq. (2) describes a linear relationship between the traveltimes and the 2-D slowness field. In principle, the slowness vector $\mathbf{u}$ may be obtained by inverting the system of equations (2). In practice, it is generally not possible to determine $\mathbf{u}$ unambiguously without introducing *a priori* information in the form of smoothing and/or damping constraints:

$$\begin{bmatrix} \mathbf{t} \\ \mathbf{0} \\ \mathbf{u}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{L} \\ \mathbf{A} \\ \mathbf{I} \end{bmatrix} \mathbf{u}\,, \tag{3}$$

where $\mathbf{A}$ is a smoothing matrix (Constable *et al.* 1987), $\mathbf{u}_0$ is a vector of damping constraints (Marquardt 1970) and $\mathbf{I}$ is the identity matrix. Eq. (3) can be written in a more compact form as

$$\mathbf{d} = \mathbf{G}\mathbf{u}. \tag{4}$$

The smoothing and damping constraints cause the system of equations (4) to be overdetermined. Because the values of $\mathbf{L}$ depend on the unknown slowness field $\mathbf{u}$, the inversion problem is non-linear. Consequently, eq. (4) must be solved iteratively (e.g. Menke 1989).

3 CONTINUOUS L_2-NORM MINIMIZATION

Algorithms that employ 'L_2-norm minimization' attempt to minimize the squared sum of the prediction error

$$\sum_{i=1}^{N} \sum_{j=1}^{M} (G_{ij} u_j - d_i)^2, \tag{5}$$

where N is the number of traveltimes plus the additional constraints (see eq. 4). There are several options for solving the classical least-squares problem. Popular choices include accumulation of the normal equations and inverting the resultant Hessian matrix and singular-value decomposition of $\mathbf{G}$ (e.g. Menke 1989). For very large data sets, the conjugate gradient methods (e.g. LSQR) of Paige & Saunders (1982) prove to be particularly efficient. All of these methods involve a directed search in the model space. They can only be applied when the model parameter range is continuous.

4 CONTINUOUS L_1 NORM MINIMIZATION

Algorithms that employ 'L_1-norm minimization' attempt to minimize the absolute difference of the prediction error

$$\sum_{i=1}^{N}\sum_{j=1}^{M}|G_{ij}u_j - d_i|. \tag{6}$$

Linear programming is typically used for this purpose (Dantzig 1963; Menke 1989; Press *et al.* 1992). The overdetermined system of equations (4) must be converted into an appropriate form for L_1-norm minimization. Following Menke (1989), we set

$$\begin{aligned}
&\sum_{j=1}^{M} G_{ij}u_j + x_i - \alpha_i = d_i \\
&\sum_{j=1}^{M} G_{ij}u_j - x_i' + \alpha_i = d_i \\
&u_i \geq 0 \qquad \alpha_i \geq 0 \\
&x_i \geq 0 \qquad x_i' \geq 0.
\end{aligned} \tag{7}$$

Through linear programming, we attempt to minimize the objective function $\sum_{i=1}^{N}\alpha_i$. Note, that $x_i - \alpha_i$ and $-x_i' + \alpha_i$ represent the errors in fitting eqs (7). Since they are differences between positive quantities, these errors can be positive or negative. If $\sum_{j=1}^{M} G_{ij}u_j - d_i$ is positive, the first set of constraints requires that $\alpha_i \geq \sum_{j=1}^{M} G_{ij}u_j - d_i$, since x_i cannot be negative. The second set of constraints can then always be satisfied by choosing appropriate positive values of x_i'. In comparison, if $\sum_{j=1}^{M} G_{ij}u_j - d_i$ is negative, then the first set of constraints can always be satisfied by choosing any appropriate x_i, but then the second set of constraints requires that $\alpha_i \geq -\sum_{j=1}^{M} G_{ij}u_j + d_i$. Taken together, these two sets of constraints require that $\alpha_i \geq |\sum_{j=1}^{M} G_{ij}u_j - d_i|$. Minimizing $\sum_{i=1}^{N}\alpha_i$ is, therefore, equivalent to minimizing $\sum_{i=1}^{N}\sum_{j=1}^{M}|G_{ij}u_j - d_i|$. In this form, the equations can be solved directly over a continuous range using the simplex algorithm (Press *et al.* 1992).

5 DISCRETE TOMOGRAPHY

When L_1-norm minimization is combined with linear programming, a variety of options for constraining the elements of **u** can be implemented. In particular, their application in discrete tomography is facilitated by the ability to add inequality constraints, such that values of **u** are restricted to a few narrow ranges. For the case of two ranges of values, this can be achieved by extending the system of equations (7) with the following template for all M elements of **u**:

$$\begin{aligned}
&y_j^1 + y_j^2 = 1 \\
&-a_{\mathrm{L}}^1 y_j^1 + z_j^1 \geq 0 \\
&-a_{\mathrm{U}}^1 y_j^1 + z_j^1 \leq 0 \\
&-a_{\mathrm{L}}^2 y_j^2 + z_j^2 \geq 0 \\
&-a_{\mathrm{U}}^2 y_j^2 + z_j^2 \leq 0 \\
&-u_j + z_j^1 + z_j^2 = 0,
\end{aligned} \tag{8}$$

where y_j^1 and y_j^2 are dummy binary variables that can be either 0 or 1, z_j^1 and z_j^2 are dummy continuous variables, and $[a_{\mathrm{L}}^1, a_{\mathrm{U}}^1]$ and $[a_{\mathrm{L}}^2, a_{\mathrm{U}}^2]$ define the two ranges of values (L, lower bound and U, upper bound). For the case of three ranges of values, two further equations would need to be added for each element of **u**.

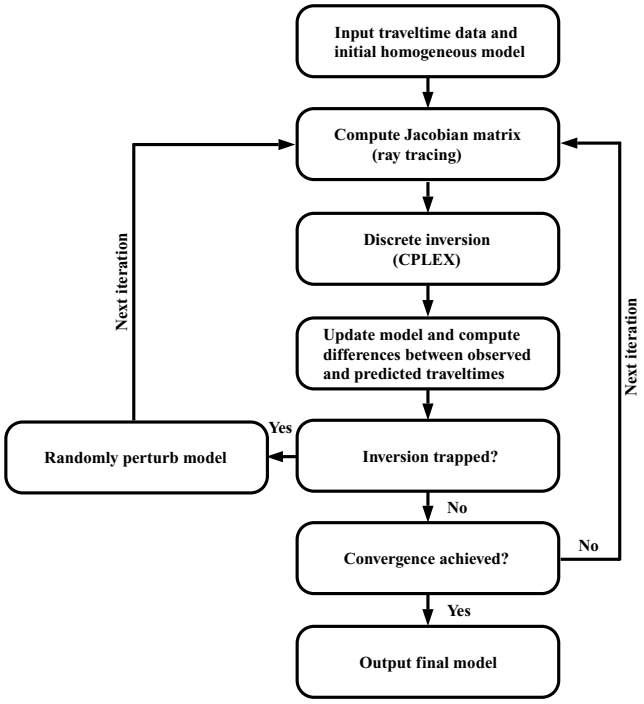

Figure 1. Flow diagram describing our discrete inversion procedure.

The discrete aspect of the solution is introduced through the binary variables y_j^1 and y_j^2. By definition, such a formulation falls into the category of mixed-integer linear programming (MILP) problems (Floudas 1995), because the objective function and the constraints are linear and some variables are integers (binary in our case). The computational time to solve such problems grows exponentially with size.

Treating the inversion as a MILP problem requires linear programming software with the added capability of handling integer values. We have used ILOG CPLEX 7.0 (ILOG 2000), which includes a branch-and-cut algorithm (Land & Doig 1960; Grötschel & Holland 1991; Padberg & Rinaldi 1991; Floudas 1995). Although it is beyond the scope of this paper to discuss details of branch-and-cut algorithms in detail, the Appendix includes a simple example that helps to explain the basic idea of the method.

For traveltime tomography, the MILP algorithm has to be applied in an iterative fashion (Fig. 1). Starting with an input homogeneous velocity model based on the estimated velocity of the host medium and straight rays, the MILP algorithm attempts to minimize the objective function subject to the constraints described in eqs (7) and (8). The resulting solution of velocity values is used to recompute the rays (now curved) and the MILP algorithm is restarted. This scheme is repeated until a convergence criterion is reached (i.e. when the data misfit is below a predefined value, determined on the basis of the traveltime pick accuracy). If during the course of the inversion process a solution occurs twice in a row, a random perturbation is made to the model to allow the inversion to escape from insignificant local minima and to continue. In our algorithm, we simply set the velocity of a certain number (e.g. 30 per cent) of randomly chosen cells back to the value of the input homogeneous model. Note, that this latter step merely causes the inversion program to explore other parts of the model space. It does not guarantee success. Nevertheless, it was found to be a dependable process for all tested models. To ensure that a global minimum is found, the entire inversion process should be repeated several times (see Section 7.1.4).

6 JOINT DISCRETE TOMOGRAPHY

Our discrete tomography formulation for a single data set can be readily extended to joint inversions of two or more data sets that are sensitive to different physical parameters. For example, in searching for air-filled cavities using the seismic and ground-penetrating radar (georadar) methods, the known seismic and georadar velocities in air are ∼300 m s^{-1} and ∼0.3 m ns^{-1}, respectively. Since the seismic and georadar velocities of the host rock may be of the order of 3000 m s^{-1} and 0.1 m ns^{-1}, respectively, seismic waves experience major decreases in velocity at rock–air interfaces, whereas georadar waves experience major increases. Contrasts of opposite sign occur at air–rock interfaces. Yet the shape of the cavity appears practically identical for both methods (assuming wavelengths of the two wave types are comparable), a property that is exploited in our algorithm.

The approach outlined below is formulated for the cavity detection problem using a combination of the seismic and georadar methods, for which only two discrete physical properties are relevant. Extensions to other methods and additional physical properties are straightforward.

As a first step, the two systems of equations for the individual inversion problems are incorporated into a single combined system:

$$\begin{bmatrix} \mathbf{G}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_2 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \tag{9}$$

where the indices 1 and 2 refer to seismic and georadar parameters, respectively. The model discretization for the slownesses must be identical for both types of data, but the distribution of sources and receivers may be different. In a second step, the combined system of equations is transformed into a linear programming form according to eqs (7). Finally, the following template of equations, which is analogous to that introduced for single discrete inversions in eqs (8), is added for all M pairs of slownesses

$$\begin{aligned}
&y_j^1 + y_j^2 = 1 \\
&-a_{\mathrm{L}}^1 y_j^1 + z_j^1 \geq 0 \\
&-a_{\mathrm{U}}^1 y_j^1 + z_j^1 \leq 0 \\
&-a_{\mathrm{L}}^2 y_j^2 + z_j^2 \geq 0 \\
&-a_{\mathrm{U}}^2 y_j^2 + z_j^2 \leq 0 \\
&-b_{\mathrm{L}}^1 y_j^1 + z_j^3 \geq 0 \\
&-b_{\mathrm{U}}^1 y_j^1 + z_j^3 \leq 0 \\
&-b_{\mathrm{L}}^2 y_j^2 + z_j^4 \geq 0 \\
&-b_{\mathrm{U}}^2 y_j^2 + z_j^4 \leq 0 \\
&-u_j + z_j^1 + z_j^2 = 0 \\
&-u_{j+M} + z_j^3 + z_j^4 = 0,
\end{aligned} \tag{10}$$

where y_j^1 and y_j^2 are dummy binary variables, z_j^1, z_j^2, z_j^3 and z_j^4 are dummy continuous variables, and $[a_{\mathrm{L}}^1, a_{\mathrm{U}}^1]$, $[a_{\mathrm{L}}^2, a_{\mathrm{U}}^2]$, $[b_{\mathrm{L}}^1, b_{\mathrm{U}}^1]$ and $[b_{\mathrm{L}}^2, b_{\mathrm{U}}^2]$ define the two ranges of values for the seismic (a) and georadar (b) slowness values, respectively. For the case of three discrete ranges of values, four additional equations are required. The slowness vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are merged into a single vector, where u_j ($j = 1, \ldots, M$) are the seismic slownesses and u_j ($j = M + 1, \ldots, 2M$) are the georadar slownesses.

7 SYNTHETIC TESTS WITH HIGH VELOCITY CONTRASTS

7.1 A single body embedded in quasi-homogeneous media

To test the MILP inversion algorithms, a generic 20 × 20 unit model with a single 6 × 6 unit inclusion in the centre was created. Initially, two versions of the model were considered: a low-velocity inclusion within a high-velocity homogeneous medium (Fig. 2a; velocity ratio of 4/6) and a high-velocity inclusion within a low-velocity homogeneous medium (Fig. 2b; velocity ratio of 6/4). A combination of these two models could represent an ice lens within a crystalline host rock, such that the ice lens appears as a low-velocity anomaly to seismic waves (e.g. 3500 versus 5300 m s^{-1}) and a high-velocity anomaly to georadar waves (e.g. 0.17 versus 0.11 m ns^{-1}). Random velocity fluctuations of ≤5 per cent superimposed on both models introduced traveltime fluctuations of ∼1 per cent. Using 11 sources along the left-hand edge and 11 receivers along the right-hand edge of the model, asymptotic ray theory was used to generate synthetic traveltimes for both types of wave. To highlight the effects of the two velocity anomalies, relative reduced traveltimes (RRT) were computed as follows:

$$\mathrm{RRT} = 100 \frac{t^{\mathrm{calc}} - t^{\mathrm{homo}}}{t^{\mathrm{calc}}}, \tag{11}$$

where t^{calc} were the traveltimes calculated using the models in Figs 2(a) and (b) and t^{homo} were the traveltimes calculated using models with either a velocity of 4 (Fig. 2a) or a velocity of 6 (Fig. 2b). The RRTs for each model are shown in Figs 2(c) and (d), and the ray paths for all source–receiver pairs are plotted in Figs 2(e) and (f).

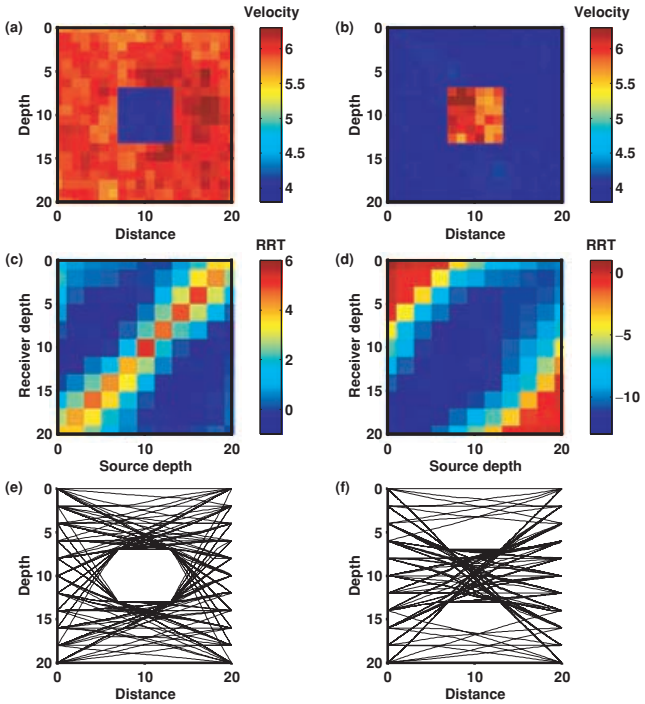

Figure 2. Velocity models comprising constant-velocity media with superimposed random fluctuations (≤5 per cent; note that the random fluctuations are difficult to see in the dark blue regions) and (a) an embedded low-velocity body and (b) an embedded high-velocity body. Parts (c) and (d) show relative reduced traveltimes (RRT = 100($t^{\mathrm{calc}} - t^{\mathrm{homo}}$)/$t^{\mathrm{calc}}$) for velocity models in (a) and (b), respectively. Parts (e) and (f) display ray path distributions for all transmitters to all receivers for models (a) and (b), respectively.

The mostly positive RRTs of Fig. 2(c) are caused by the rays that curve around the low-velocity body. In contrast, the mostly negative RRTs of Fig. 2(d) are caused by rays being channelled through the high-velocity body of Fig. 2(d). It is noteworthy that the RRTs associated with the low-velocity body lie between −1 and 6 per cent (Fig. 2c), whereas those associated with the high-velocity body are as negative as −13 per cent (Fig. 2d). The ray distribution in Fig. 2(e) suggests that the rough location and shape of the low-velocity body may be delineated in an inversion process, but the actual anomalous velocity values are unlikely to be resolved. The concentration of rays through the high-velocity body (Fig. 2f) causes gaps above and below the anomaly that may introduce resolution problems at these locations.

All inversions were carried out with a cell size of 1 × 1 unit. The smoothing and damping constraints were chosen by trial and error and the initial models were homogeneous with the velocity set to that of the host medium. Since the solution space in discrete tomography is restricted to two or three *a priori* known ranges of values, it is appropriate to initiate the inversion with the velocity of the host medium, or an approximation thereof.

7.1.1 *Conventional velocity tomograms*

The low-velocity body is detected on the conventional (least-squares) tomogram, but its aspect ratio is distorted and the anomalous velocities are strongly overestimated (5.1–5.4 versus 4.0; Fig. 3a). Even with these moderately low velocities, the fastest rays are those that circumvent the anomalous body. Note how few rays enter the low-velocity body in Fig. 3(e), with only one ray traversing it. The high-velocity anomaly is somewhat better resolved than the low-velocity one, but the velocity contrast is underestimated (4.8–5.1 versus 6.0; Fig. 3b). For both cases, the slightly asymmetrical tomograms are largely the result of the ray-tracing program and the small random traveltime fluctuations.

There is a good match between the RRTs in Figs 3(c) and (d) and between the RRTs in Figs 2(c) and (d), indicating that the inverted models explain the data well. The degree of agreement between true t^{true} and predicted t^{pre} traveltimes is further quantified by the average relative residual (ARR) defined as

$$\text{ARR} = \frac{100}{N}\sum_{i=1}^{N}\frac{|t_i^{\text{true}} - t_i^{\text{pre}}|}{t_i^{\text{true}}}. \tag{12}$$

As indicated in Figs 3(a) and (b), the ARRs (0.25 and 0.34 per cent) are well below the 1 per cent traveltime variations caused by the random velocity fluctuations.

The ray diagrams in Figs 3(e) and (f) are similar to those in Figs 2(e) and (f), except the smoothing constraints enable a few rays to enter the low-velocity body (Fig. 3e) and travel through the regions above and below the high-velocity body (Fig. 3f). These same smoothing constraints cause the anomalous bodies to appear somewhat blurred in Figs 3(a) and (b).

7.1.2 *Discrete velocity tomograms*

For the discrete tomography, it is assumed that the velocities of the host material and anomalous bodies are approximately known; velocities are forced to lie between either 3.8 and 4.2 or 5.7 and 6.3. The resultant tomogram of Fig. 4(a) shows the correct location

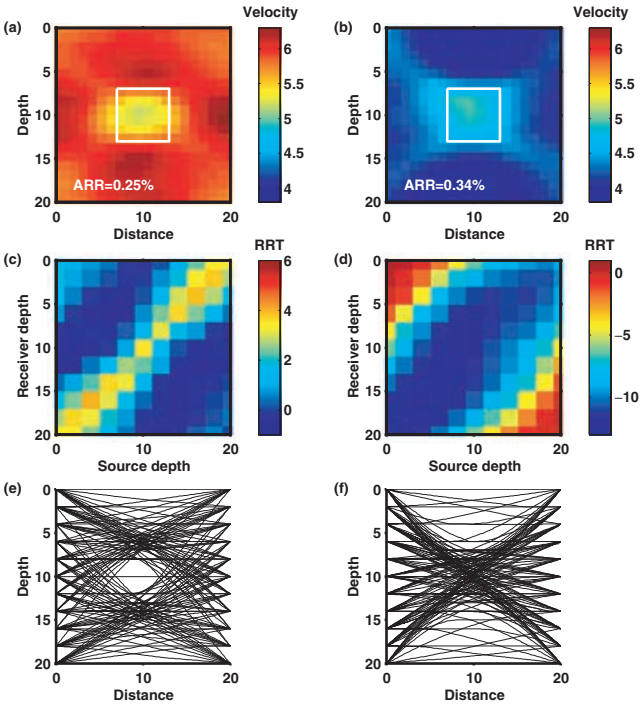

Figure 3. Parts (a) and (b) are conventional velocity tomograms determined from the two suites of traveltimes computed for velocity models in Figs 2(a) and (b), respectively. Solid lines outline the true boundaries of the embedded bodies. The average relative residual (ARR= $\frac{100}{N}\sum_{i=1}^{N}|t_i^{\text{true}} - t_i^{\text{pre}}|/t_i^{\text{true}}$) is shown for both tomograms. Parts (c) and (d) show relative reduced traveltimes (RRT = $100(t^{\text{calc}} - t^{\text{homo}})/t^{\text{calc}}$) for velocity models in (a) and (b), respectively. Parts (e) and (f) display ray path distributions for all transmitters to all receivers for models (a) and (b), respectively.

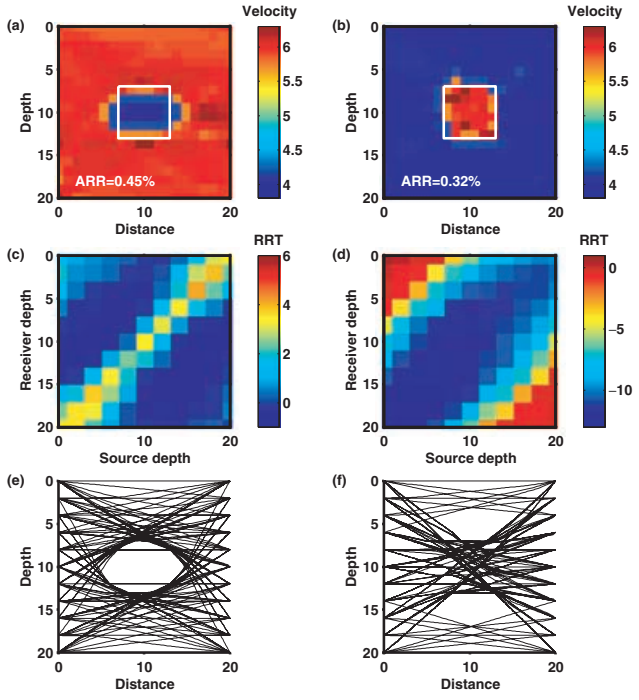

Figure 4. Parts (a) and (b) are discrete velocity tomograms determined by individually inverting the two suites of traveltimes computed for velocity models in Figs 2(a) and (b), respectively. Solid lines outline the true boundaries of the embedded bodies. The average relative residual (ARR = $\frac{100}{N}\sum_{i=1}^{N}|t_i^{\text{true}} - t_i^{\text{pre}}|/t_i^{\text{true}}$) is shown for both tomograms. Parts (c) and (d) show relative reduced traveltimes (RRT = $100(t^{\text{calc}} - t^{\text{homo}})/t^{\text{calc}}$) for velocity models in (a) and (b), respectively. Parts (e) and (f) display ray path distributions for all transmitters to all receivers for models (a) and (b), respectively.

and velocity of the low-velocity body, but its shape is distorted in a similar fashion to that of the conventional tomogram of Fig. 3(a). Features of the high-velocity body are quite well resolved (Fig. 4b). The RRT patterns of Figs 4(c) and (d) are close to those of Figs 2(c) and (d), respectively, and the ARRs (0.45 and 0.32 per cent) are again well below the random traveltime fluctuations. Since the velocities are restricted to lie within two narrow ranges about the true values, the ray paths in Figs 4(e) and (f) are close to the true ray paths in Figs 2(e) and (f).

7.1.3 Joint discrete velocity tomograms

Excellent reconstructions of both the low- and high-velocity bodies are achieved using the discrete joint inversion algorithm (Figs 5a and b). Considering the good fits, it is not surprising that the RRT patterns of Figs 2(c) and (d) are well produced in Figs 5(c) and (d), and the ARRs (0.15 and 0.30 per cent) are again well below the random traveltime fluctuations. The improved results are not only a result of the effective increase in data (242 instead of 121 traveltimes), but they are also the result of an improved ray distribution (Fig. 5e); all parts of the model are now reasonably well covered by crossing rays. Note, that simply adding more sources and receivers along the same length of boreholes would not substantially improve the tomograms shown in Figs 3(a) and 4(a).

7.1.4 Convergence characteristics

Although the discrete tomograms (Figs 4a,b and 5a,b) are superior to the conventional ones (Figs 3a and b), convergence of the

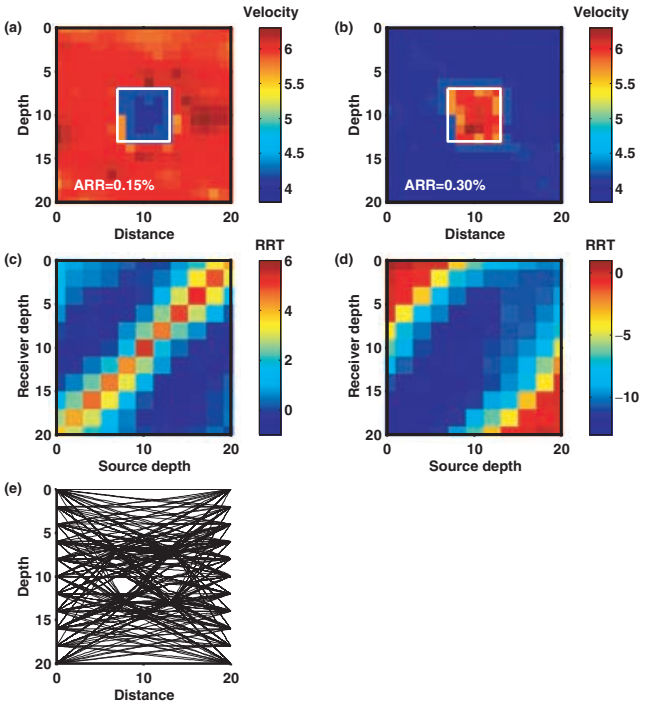

Figure 5. Parts (a) and (b) are discrete velocity tomograms determined by jointly inverting the two suites of traveltimes computed for velocity models for Figs 2(a) and (b), respectively. Solid lines outline the true boundaries of the embedded bodies. The average relative residual (ARR= $\frac{100}{N}\sum_{i=1}^{N}|t_i^{\text{true}}-t_i^{\text{pre}}|/t_i^{\text{true}}$) is shown for both tomograms. Parts (c) and (d) show relative reduced traveltimes (RRT $=100(t^{\text{calc}}-t^{\text{homo}})/t^{\text{calc}}$]) for velocity models in (a) and (b), respectively. (e) Combined ray path distribution for velocity tomograms of (a) and (b).

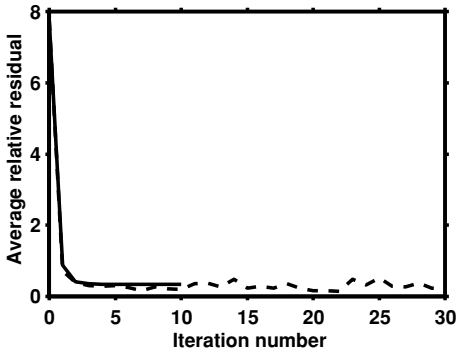

Figure 6. Average relative residuals (ARR $=\frac{100}{N}\sum_{i=1}^{N}|t_i^{\text{true}}-t_i^{\text{pre}}|/t_i^{\text{true}}$) as functions of the iteration number for conventional tomography (solid line) and discrete tomography (dashed line) applied to traveltime data computed for the velocity model in Fig. 2(b). The average relative residual at 'iteration 0' was calculated for the input homogeneous model.

MILP algorithm is more erratic than that of the least-squares approach. Fig. 6 shows the ARRs as functions of iteration number for the conventional and single discrete inversion runs that resulted in Figs 3(b) and 4(b). The form of the ARR function for the least-squares inversion is fairly standard, with a rapid decay followed by gradually decreasing values. In contrast, the equivalent curve for the discrete inversion falls rapidly to a minimum, then oscillates over $\sim$20 iterations before falling to reach a slightly lower minimum. The numerous local minima are probably caused by non-linear effects introduced by the binary variables: in continuous inversions, the velocities are modified gradually from one iteration to the next, such that ray paths change gradually, whereas in discrete inversions, large velocity contrasts are created in the model domain, such that ray paths may change abruptly from one iteration to the next. For large velocity contrasts, many iterations may be required before an acceptable data misfit is obtained.

Using the model of Fig. 2(b) as an example, we now wish to explore the robustness of our solutions. Least-squares methods are known to be fairly robust when applied to high-quality dense and well-distributed data that contain moderate amounts of random noise. They converge to similar solutions from different starting models. Moreover, there exist a number of means to test the reliability of conventional tomograms (e.g. resolution matrices and singular-value spectra; Menke 1989). In comparison, it is unclear how the non-linear influence associated with restricting some variables to be binary affects the robustness of the MILP results. A simple method to test the dependability of MILP-based models is to repeat the inversions several times. Since MILP algorithms include random components, the resulting models should scatter about the most probable solutions in the model space. Fig. 7 shows discrete velocity tomograms resulting from runs with the same suite of traveltimes generated for the velocity model of Fig. 2(b), but with different starting models (the velocities of the initial homogeneous models varied by ±5 per cent) and different random 'seeds'. Differences between the tomograms are a measure of the dependability of the reconstructions. For this and other examples shown in this paper, very similar solutions and ARRs are obtained, regardless of the random perturbations. As a rule, this testing procedure (i.e. several runs of the discrete tomography program) should be followed for the inversion of all data sets.

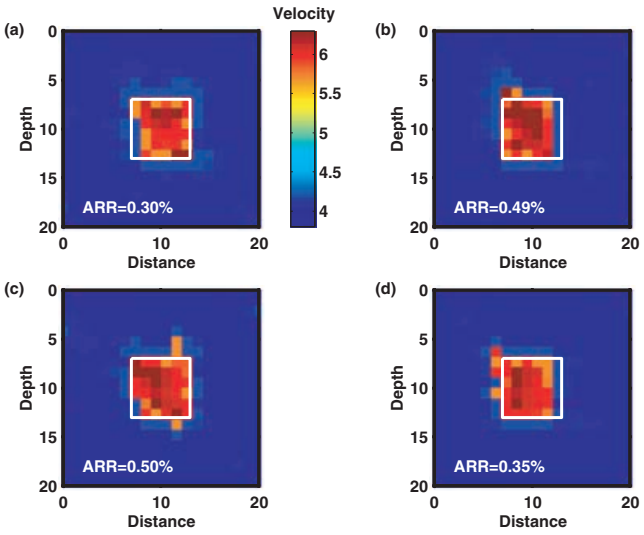

Figure 7. (a)–(d) Discrete velocity tomograms resulting from runs with the same suite of traveltimes computed for the velocity model of Fig. 2(b), but with different starting models and different random 'seeds' (see the text).

7.2 Two separate bodies of the same material type embedded in quasi-homogeneous media

To gain further insight into the behaviour of our discrete tomography algorithm, additional tests with more complex models have been performed. As for the first example (Figs 2–5), stochastic velocity fluctuations of $\leq$5 per cent were superimposed on the basic model values (e.g. two bodies with anomalous velocities embedded in homogeneous media).

Results for models with two embedded inclusions of the same type are shown in Fig. 8. In the conventional tomograms, the two square low-velocity bodies appear as horizontally stretched structures with velocities that are too high (5.5–5.7 versus 4.0; Fig. 8c), whereas the two square high-velocity bodies are smeared diagonally with velocities that are too low (4.8–5.3 versus 6.0; Fig. 8d). The discrete tomograms identify two separate bodies with the correct velocities (Figs 8e and f), but their shapes are not well reproduced. In comparison, the joint discrete inversion produces tomograms that reveal nearly square bodies at their correct positions (Figs 8g and h). In all cases, the ARRs are below the 1 per cent random traveltime fluctuations. This indicates that all models are numerically equivalent; improvements are only possible by increasing the source and receiver apertures.

7.3 Two separate bodies of distinct material types embedded in quasi-homogeneous media

For the next suite of tests, three discrete velocity values were used. In Fig. 9(a), high- and low-velocity bodies were embedded in a homogeneous medium of velocity 4.6. Again, stochastic velocity fluctuations of $\leq$5 per cent were superimposed on the basic model values. In Fig. 9(b), the bodies were interchanged and different random fluctuations superimposed; because the sources and receivers are symmetric, Figs 9(a) and (b) show two realizations of essentially the same model.

In the conventional tomograms of Figs 9(c) and (d), the anomalous bodies are barely recognizable, with velocities that are too low (5.1–5.3 versus 6.0) or too high (4.4–4.6 versus 3.8). The discrete tomograms reveal two separate bodies with the correct velocities

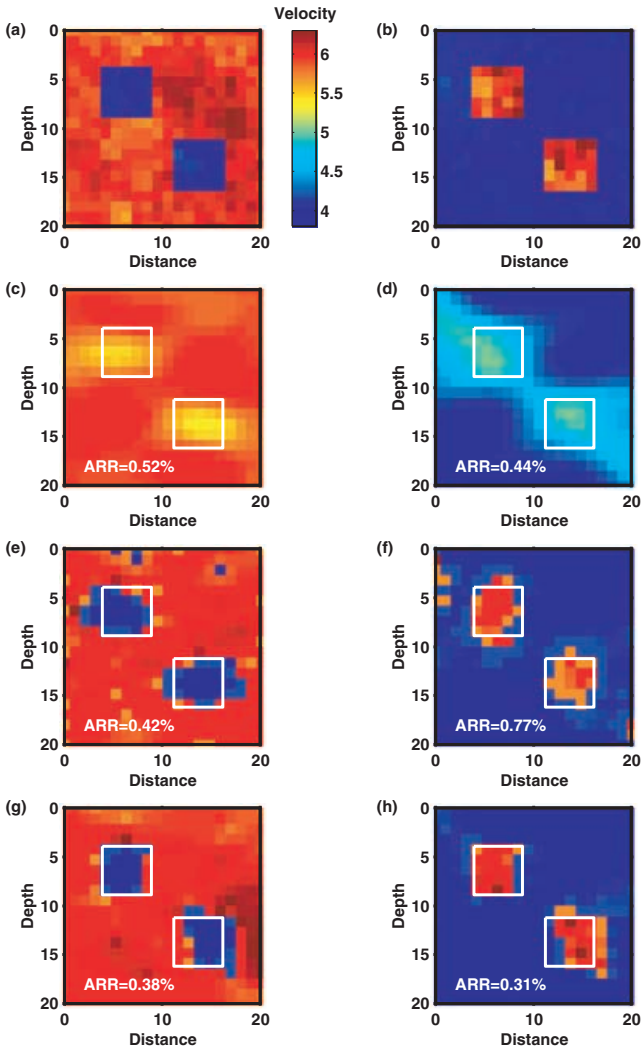

Figure 8. Velocity models comprising constant-velocity media with superimposed random fluctuations ($\leq$5 per cent; note that the random fluctuations are difficult to see in the dark blue regions) and (a) two embedded low-velocity anomalies and (b) two embedded high-velocity anomalies. Parts (c) and (d) are conventional velocity tomograms determined from the two suites of traveltimes generated by ray tracing through the velocity models shown in (a) and (b), respectively. Parts (e) and (f) are discrete velocity tomograms determined by individually inverting the two suites of traveltimes. Parts (g) and (h) are discrete velocity tomograms determined by jointly inverting the two suites of traveltimes. Average relative residuals (ARR $= \frac{100}{N}\sum_{i=1}^{N}|t_i^{\rm true}-t_i^{\rm pre}|/t_i^{\rm true}$) and boundaries of true velocity anomalies are shown on each tomogram.

(Figs 9e and f), but their shapes and positions are incorrect. Note how the low-velocity bodies in both discrete tomograms are somewhat horizontally elongated. In contrast, the joint discrete inversion produces tomograms that show nearly square bodies at their correct locations (Figs 9g and h) and lower ARRs.

7.4 Two connected bodies of distinct material types embedded in quasi-homogeneous media

The fourth suite of tests is similar to the third one (homogenous medium of velocity 4.6 with embedded high- (6.0) and low-velocity (3.8) bodies and superimposed stochastic velocity fluctuations of $\leq$5 per cent), except the anomalous bodies are smaller and joined

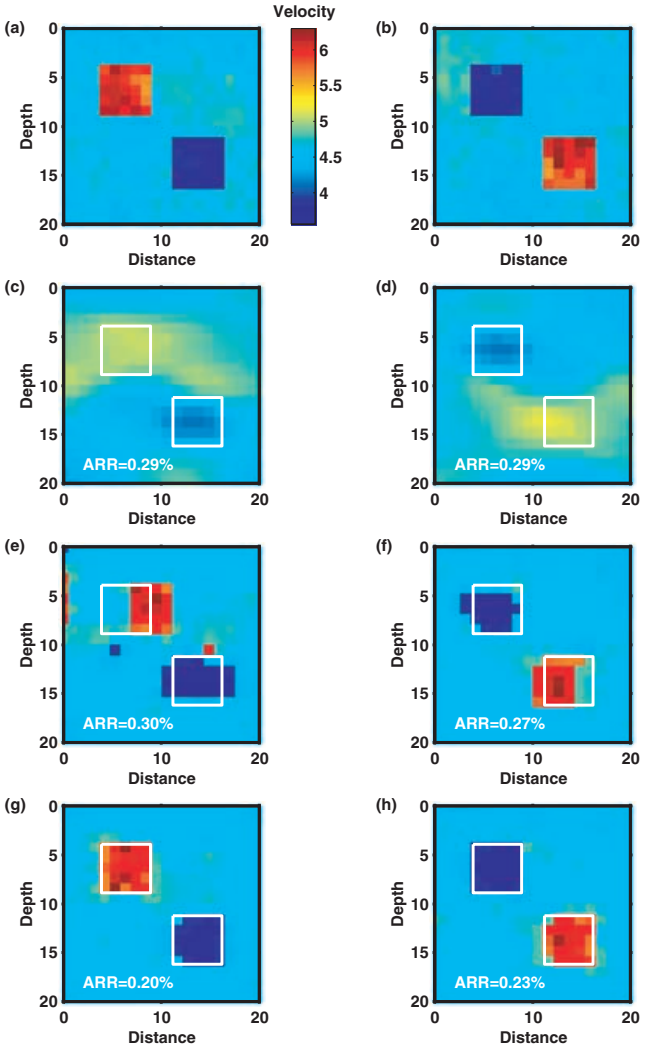

Figure 9. Velocity models comprising constant-velocity media with superimposed random fluctuations (≤5 per cent; different for each model; note that the random fluctuations are difficult to see in the dark blue regions) and symmetric (a) high- and low-velocity anomalies and (b) low- and high-velocity anomalies. Parts (c) and (d) are conventional velocity tomograms determined from the two suites of traveltimes generated by ray tracing through the velocity models shown in (a) and (b), respectively. Parts (e) and (f) are discrete velocity tomograms determined by individually inverting the two suites of traveltimes. Parts (g) and (h) are discrete velocity tomograms determined by jointly inverting the two suites of traveltimes. Average relative residuals ($\mathrm{ARR} = \frac{100}{N}\sum_{i=1}^{N}|t_i^{\mathrm{true}} - t_i^{\mathrm{pre}}|/t_i^{\mathrm{true}}$) and boundaries of true velocity anomalies are shown on each tomogram.

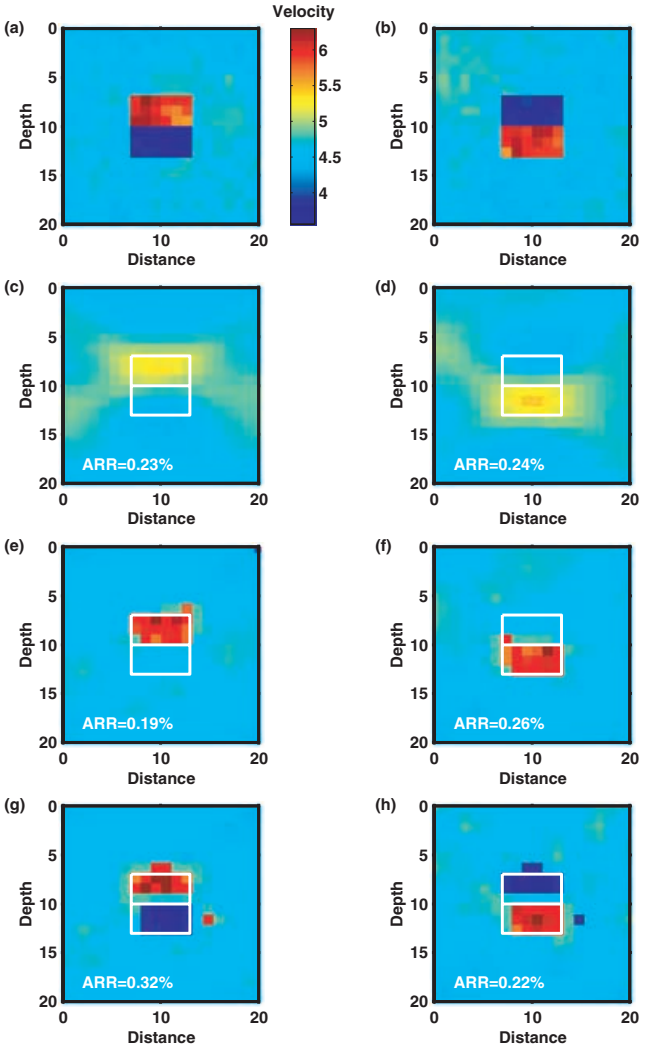

Figure 10. Velocity models comprising constant-velocity media with superimposed random fluctuations (≤5 per cent; different for each model; note that the random fluctuations are difficult to see in the dark blue regions) and (a) connected high- and low-velocity anomalies and (b) connected low- and high-velocity anomalies. Parts (c) and (d) are conventional velocity tomograms determined from the two suites of traveltimes generated by ray tracing through the velocity models shown in (a) and (b), respectively. Parts (e) and (f) are discrete velocity tomograms determined by individually inverting the two suites of traveltimes. Parts (g) and (h) are discrete velocity tomograms determined by jointly inverting the two suites of traveltimes. Average relative residuals ($\mathrm{ARR} = \frac{100}{N}\sum_{i=1}^{N}|t_i^{\mathrm{true}} - t_i^{\mathrm{pre}}|/t_i^{\mathrm{true}}$) and boundaries of true velocity anomalies are shown on each tomogram.

together (Figs 10a and b; note, we are again showing two realizations of essentially the same model).

As for previous results, the conventional tomograms are characterized by symmetric patterns. The low-velocity bodies are just barely distinguishable and the velocities of the high-velocity bodies are underestimated (Figs 10c and d; 5.1–5.3 versus 6.0). Considering the strength of the artefacts (e.g. the smeared high-velocity values on either side of the high-velocity body), it is unlikely that the low-velocity body would be identified as significant in a field data set. Lack of resolution in the low-velocity regions is caused by focusing of rays through the high-velocity bodies. Since this type of ray coverage imposes inherent limitations on resolution, it is not surprising that the discrete inversions fail to detect the low-velocity body. The individual discrete tomograms only reveal the correctly reconstructed high-velocity bodies (Figs 10e and f). For this recording configuration, a low-velocity body adjacent to a high-velocity body is effectively invisible. As a result of the complementary nature of the two traveltime data sets, the joint discrete inversion reconstructs both bodies approximately correctly (Figs 10g and h). Similarly low ARRs (0.19–0.32 per cent) are obtained for all inversions.

8 CAVITY DETECTION WITH SEISMIC AND GEORADAR METHODS

To test the discrete inversion algorithms on more realistic data, appropriate finite-difference algorithms (Robertsson *et al.* 1994;

Holliger & Bergmann 2002) are used to generate synthetic seismograms and radargrams for models of a rock mass containing a cavity. Velocities within the cavity are either those of air (seismic, ~300 m s^{-1}; georadar, ~0.3 m ns^{-1}) or water (seismic, ~1600 m s^{-1}; georadar, ~0.03 m ns^{-1}), whereas the velocities of the host rock are set to typical sedimentary rock values (seismic, 3300 m s^{-1}; georadar, 0.11 m ns^{-1}). Both the seismic and georadar velocity contrasts in these models are much higher (factors of 2–10) than those in the previous suites of models. Such high-velocity contrasts are particularly problematic for most conventional tomographic inversion routines.

The model dimensions are 20 × 20 m^2 with a cell size of 1 m. Stochastic velocity fluctuations of ≤5 per cent are, again, added to the host medium. The source–receiver geometries are the same as those employed in previous examples. First-arrival traveltimes are picked from the synthetic traces. We estimate that the traveltime picks have an accuracy of approximately ±2 per cent.

8.1 An air cavity embedded in quasi-homogeneous host rock

Our first example using synthetic seismograms and radargrams is a single air-filled cavity within a host sedimentary rock (Figs 11a and b). Seismic and georadar sections for sources located at 10 m depth are presented in Figs 12(a) and (b). The seismic section is mainly characterized by first arrivals that circumvent the cavity. They have highly variable amplitudes with the weakest signals recorded at receivers in the shadow of the cavity. The georadar section contains several prominent phases: first arrivals that pass once through the cavity, and strong secondary arrivals that comprise direct and diffracted waves that travel entirely within the host material (dominant secondary phases recorded at the upper and lower receivers) and waves that are doubly reflected within the cavity (i.e. a form of multiple; dominant secondary phases recorded at the central receivers).

Conventional inversions perform rather poorly (Figs 11c and d), probably because of the large velocity contrasts involved. The cavity is barely visible in the seismic tomogram (Fig. 11c) and although the georadar tomogram includes a high-velocity anomaly at the correct location (Fig. 11d), the estimated velocity contrast is far too small for an air-filled cavity.

For the discrete inversions, we assume that the cavity could be either air- or water-filled. This requires consideration of three different velocities (air, water, host rock). For the model geometry shown in Fig. 11(a), a ~20 per cent negative velocity contrast is sufficient to ensure that the first-arriving energy travels entirely around the anomaly. Consequently, on the basis of the seismic data alone it is not possible to distinguish between air- and water-filled cavities. It is, therefore, not surprising that the discrete inversion of the seismic data suggests the presence of a water-filled cavity (Fig. 11e). In comparison, the georadar tomogram resolves the location and shape of the cavity well and yields the correct velocity (Fig. 11f).

Significantly improved models are obtained from the joint discrete inversion (Figs 11g and h). As a result of the seismic and georadar air velocity coupling (eqs 10), the location and shape of the cavity are well reconstructed and all velocities are recovered correctly. Although all ARR values are well below the ±2 per cent traveltime reading accuracy, the ARRs for the discrete inversion of the georadar data are uniformly higher (1.00–1.19 per cent) than the others. It is also noteworthy that the ARRs are similar for the discrete seismic tomograms in Figs 11(e) and (g), but the characteristics of the resultant anomalous structures are quite different. This indicates an inherent lack of information in the seismic data.

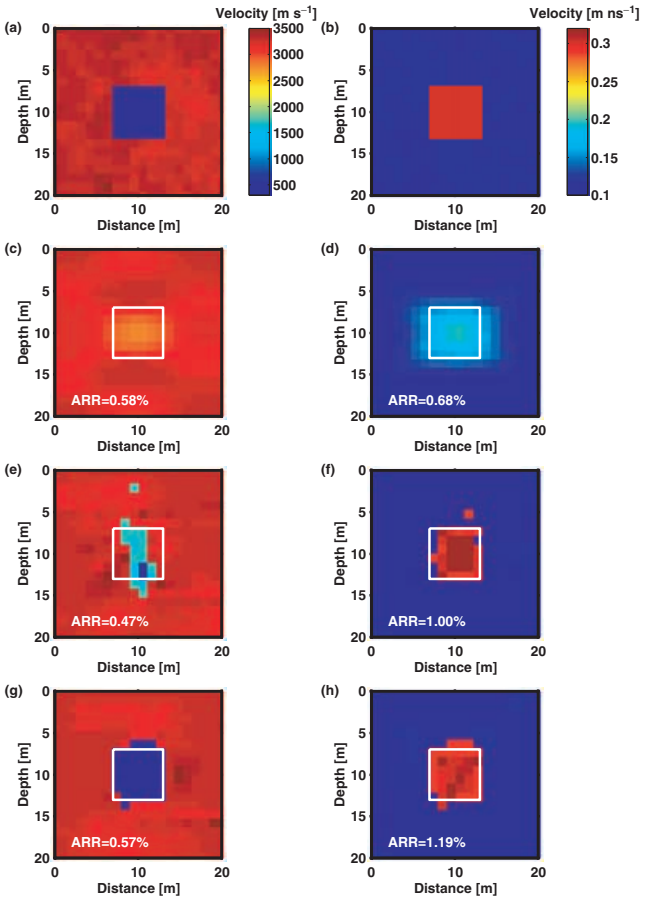

Figure 11. Seismic and georadar velocity models representing the same air-filled cavity. Velocity models comprising constant-velocity media with superimposed random fluctuations (≤5 per cent; note that the random fluctuations are difficult to see in the dark blue regions) and (a) a seismic low-velocity (300 m s^{-1}) cavity and (b) a georadar high-velocity (0.3 m ns^{-1}) cavity. Parts (c) and (d) are conventional velocity tomograms determined from the suites of traveltimes picked from finite-difference synthetic seismic and georadar sections (Fig. 12) computed for the velocity models shown in (a) and (b), respectively. Parts (e) and (f) are discrete velocity tomograms determined by individually inverting the two suites of traveltimes. Parts (g) and (h) are discrete velocity tomograms determined by jointly inverting the two suites of traveltimes. Average relative residuals ($\mathrm{ARR} = \frac{100}{N}\sum_{i=1}^{N}|t_i^{\mathrm{true}} - t_i^{\mathrm{pre}}|/t_i^{\mathrm{true}}$) and boundaries of true velocity anomalies are shown on each tomogram.

8.2 A cavity half-filled with water embedded in quasi-homogeneous host rock

The second example corresponds to a cavity half-filled with water and half-filled with air (Figs 13a and b). Seismic and georadar sections for sources located at 10 m depth are presented in Figs 14(a) and (b). Two prominent phases are present in the seismic section: first arrivals that circumvent the cavity and waves that pass through the water-filled part of the cavity. The georadar section is similar to that derived for the totally air-filled cavity (Fig. 12b); high-velocity waves travelling through the air-filled part of the cavity dominate the georadar section (Fig. 14b).

Again, the conventional inversion of the seismic data produces rather poor results (Fig. 13c). The conventional georadar tomogram

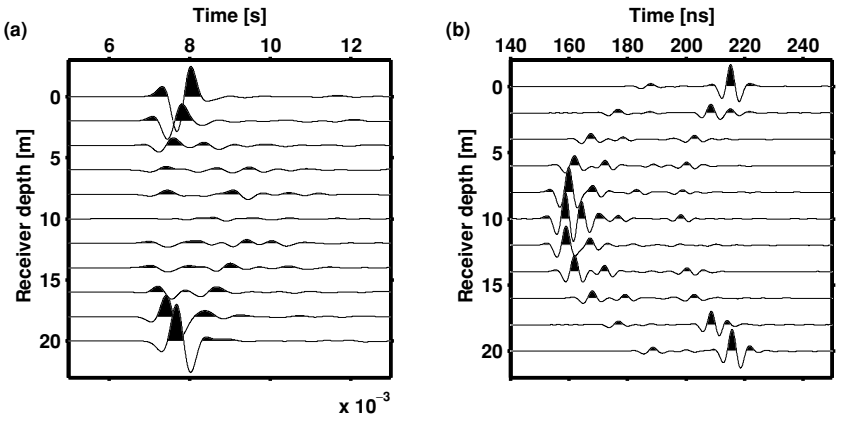

Figure 12. (a) A seismic section for the source position at 10 m depth generated from the model given in Fig. 11(a). (b) A georadar section for the source position at 10 m depth generated from the model given in Fig. 11(b). Traces are plotted with true relative amplitudes.

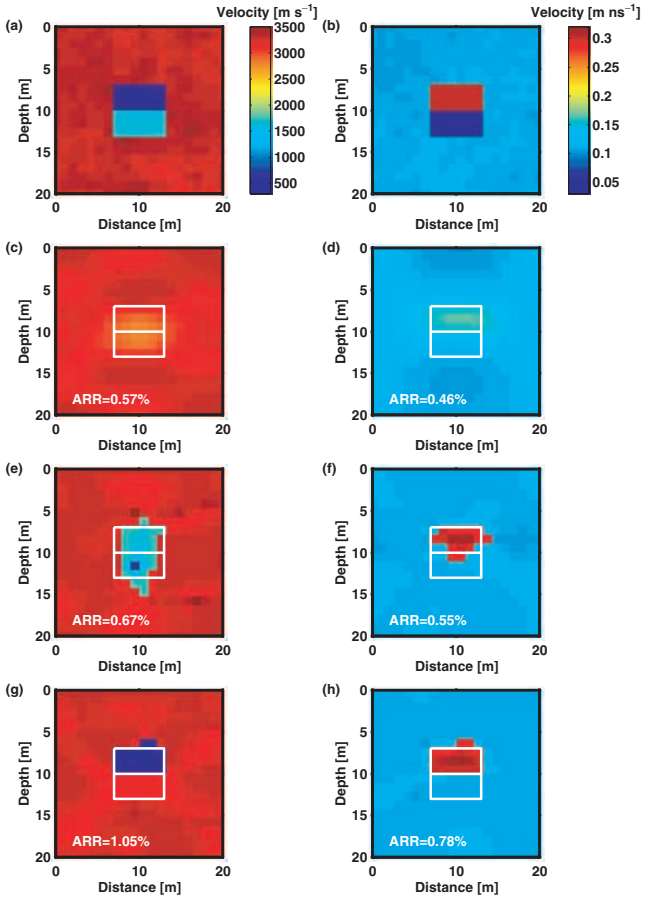

Figure 13. As for Fig. 11, except the lower half of the cavity is filled with water and the upper half is filled with air.

reveals the presence of the air-filled part of the cavity, but the estimated velocity is much too low (0.18 versus 0.3 m ns^{-1}). Considering the inadequacies of the conventional tomograms, the low ARRs (0.46–0.57 per cent) are surprising.

The discrete inversion of the seismic data yields a vertically elongated anomaly with the velocity of water (Fig. 13e), whereas the corresponding georadar tomogram resolves the air-filled part of the cavity (Fig. 13f). On the basis of these results, it would be concluded that a cavity exists. As mentioned for Fig. 11(e), the seismic data require the presence of low-velocity material, it can be water and/or air. The water-filled part of the cavity is difficult to resolve with this type of data, because rays focus in the adjacent high-velocity air-filled part of the cavity. Tomograms that result from the joint discrete inversion (Figs 13g and h) map well the air-filled part of the cavity, but they do not delineate the water-filled part. This is a result of the fact that the seismic and georadar data do not complement each other in this environment; in both models, the water-filled part of the cavity has lower velocities than the host rock.

9 DISCUSSION AND CONCLUSIONS

We have introduced a discrete tomography technique for individually or jointly inverting seismic and georadar crosshole data. The technique is applicable to a broad class of problems for which the propagation velocities are restricted to a few relatively narrow ranges of values. If sufficient *a priori* velocity information exists, the tomographic inversions should be reliable. For example, we have demonstrated that the technique works well when the average velocities are known to within ±5 per cent. Other tests indicate that convergence to correct velocities also occurs when velocity uncertainties are as large as ±10 per cent. In cases for which only poor velocity information is available, wide velocity ranges would have to be chosen. This would result in only limited advantages over conventional least-squares approaches. The new technique is unlikely to produce meaningful results if the average velocities fall outside the chosen velocity ranges.

Unlike conventional least-squares inversion methods, our discrete tomography technique does not provide a formal means of estimating ambiguity or, equivalently, of determining unequivocally whether the output model is the result of an insignificant local minimum in the model space or whether it is one of a number of very similar solutions distributed about the global minimum. To address this issue, each data set should be independently inverted several times and the resultant models compared. For all of the tests that we have performed, including many not shown here, the output models of all multiple runs were found to be very close to each other (e.g. Fig. 7).

Under a variety of conditions, the joint discrete inversions were found to be more robust than the individual discrete inversions. The complementary nature of the jointly inverted data sets allowed less ambiguous tomographic reconstructions to be achieved. This was caused by the substantially improved model constraints provided

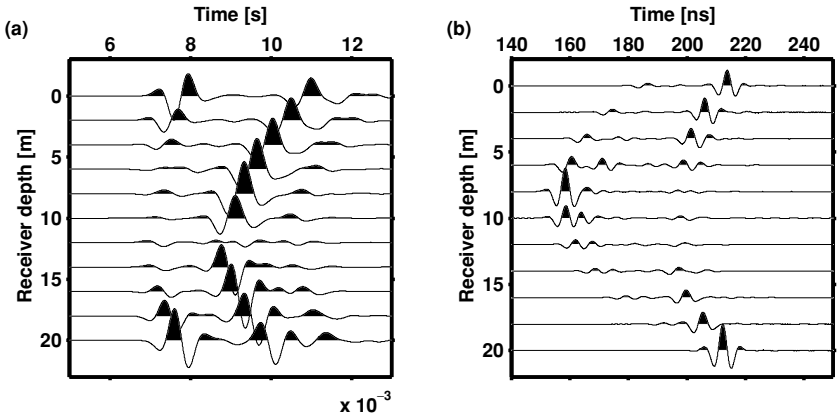

Figure 14. (a) A seismic section for a source position at 10 m depth generated from the model given in Fig. 13(a). (b) A georadar section for the source position at 10 m depth generated from the model given in Fig. 13(b). Traces are plotted with true relative amplitudes.

by the combined ray coverage of the two data sets (Figs 3–5 and 8–14). Such enhanced ray coverage for either the seismic or georadar data sets would not have been obtained by simply increasing the number of sources and receivers along the limited lengths of the boreholes, because the first-arriving energy tended to circumvent the low-velocity regions. In cases where complementary ray coverage was either very limited or not available, joint discrete inversions did not provide improved models (e.g. Fig. 13). Furthermore, despite the restrictions on the ranges of values, a certain degree of ambiguity for all of our results remained because 400 slowness cells were being derived on the basis of 121 and 242 data values in the single and joint discrete inversions, respectively.

Besides the cavity (e.g. caves, mines and tunnels) mapping problem discussed in this paper, the technique may be used for detecting ice lenses in permafrost research, ore bodies in exploration, gravel lenses in hydrogeological projects, and anthropogenic features in archaeological prospecting. Furthermore, the general technique can be extended to a wide variety of other geophysical data sets for which adequate *a priori* knowledge is available.

Compared with least-squares inversions, the MILP approach is computationally much more demanding. Typical run times for our test cases are ∼5 min for the least-squares inversions and ∼5 h for the discrete inversions. This currently limits the applicability of the method to relatively small-scale problems. However, the efficiency of linear programming algorithms is improving rapidly (for a particular problem, CPLEX 1.0 (1988) took 57 840 s and CPLEX 6.5 (1999) took 165 s) and computer performance is continuously improving.

ACKNOWLEDGMENTS

We would like to thank Klaus Holliger for providing the finite-difference programs. The project was supported financially by ETH Research Commission grant no 0-20535-98.

REFERENCES

Browne, J.A., Koshy, M. & Stanley, J.H., 1998. On the application of discrete tomography to CT-assisted engineering and design, *Int. J. Imag. Syst. Tech.,* **9,** 78–84.

Chapman, C.H. & Pratt, G.R., 1992. Traveltime tomography in anisotropic media, I. Theory, *Geophys. J. Int.,* **109,** 1–19.

Constable, S.C., Parker, R.L. & Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics,* **52,** 289–300.

Dantzig, G.B., 1963. *Linear Programming and Extensions,* Princeton University Press, Princeton.

Dyer, B.C. & Worthington, M.H., 1988. Some sources of distortion in tomographic velocity images, *Geophys. Prosp.,* **36,** 209–222.

Floudas, C.A., 1995. *Nonlinear and Mixed-Integer Optimisation,* pp. 95–107, Oxford University Press, Oxford.

Gritzmann, P., De Vries, S. & Wiegelmann, M., 2000. Approximating binary images from discrete X-rays, *SIAM J. Optim.,* **11,** 522–546.

Grötschel, M. & Holland, O., 1991. Solution of large-scale traveling salesman problems, *Math. Program.,* **51,** 141–202.

Haber, E. & Oldenburg, D., 1997. Joint inversion: a structural approach, *Inverse Problems,* **13,** 63–77.

Hering, A., Misiek, R., Gyulai, A., Ormos, T., Dobroka, M. & Dresen, L., 1995. A joint inversion algorithm to process geoelectric and surface wave seismic data, *Geophys. Prospect.,* **43,** 135–156.

Herman, G.T. & Kuba, A., 1999. Discrete tomography: foundations, algorithms and applications, Birkhäuser, Boston.

Holliger, K. & Bergmann, T., 2002. Numerical modelling of borehole georadar data, *Geophysics,* **67,** 1249–1257.

ILOG, 2000. ILOG CPLEX 7.0 Reference Manual, ILOG, Gentilly, France (http://www.ilog.com).

Jupp, D.L.B. & Vozoff, K., 1977. Resolving anisotropy in layered media by joint inversion, *Geophys. Prospect.,* **25,** 460–470.

Land, A.H. & Doig, A.G., 1960. An automatic method for solving discrete programming problems, *Econometrica,* **28,** 497–520.

Lanz, E., Maurer, H.R. & Green, A.G., 1998. Refraction tomography over a buried waste disposal site, *Geophysics,* **63,** 1414–1433.

Lines, L.R., Schultz, A.K. & Treitel, S., 1988. Cooperative inversion of geophysical data, *Geophysics,* **53,** 8–20.

Maier, D., Maurer, H.R. & Green, A.G., 1995. Joint inversion of related data sets: DC-resistivity and transient electromagnetic soundings, *1st Ann. Symp. Environ. Engin. Geophys. Soc. (European Section), Exp. Abst.* pp. 461–464.

Marquardt, D.W., 1970. Generalized inverses, ridge regression, biased linear estimation, and non-linear estimation, *Technometrics,* **12,** 591–612.

Maurer, H.R. & Green, A.G., 1997. Potential coordinate mislocations in crosshole tomography: results from the Grimsel test site, Switzerland, *Geophysics,* **62,** 1696–1709.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press, New York.

Mitchell, J.E., 2000. Branch-and-cut algorithms for combinatorial optimisation problems, *Handbook of Applied Optimisation,* Oxford University Press, Oxford.

Musil, M., Maurer, H.R., Holliger, K. & Green, A.G., 2002. Internal structure of an alpine rock glacier based on crosshole georadar traveltimes and amplitudes, *Geophysics,* submitted.

Nolet, G., 1987. *Seismic Wave Propagation and Seismic Tomography,* Reidel, Dordrecht.

Padberg, M. & Rinaldi, G., 1991. A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems, *SIAM Rev.,* **33,** 60–100.

Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw.,* **8,** 43–71.

Press, W.H., Teukolsky, S.A., Vetterrling, W.T. & Flannery, B.P., 1992. *Numerical Recipes in C,* Cambridge Univeristy Press, Cambridge.

Raiche, A.P., Jupp, D.L.B., Rutters, H. & Vozoff, K., 1985. The joint use of coincident loop transient electromagnetic and Schlumberger sounding to resolve layered structures, *Geophysics,* **50,** 1618–1627.

Robertsson, J.O.A., Blanch J.O. & Symes W.W., 1994. Viscoelastic finite-difference modeling, *Geophysics,* **59,** 1444–1456.

Sandberg, A.K., 1993. Examples of resolution improvement in geoelectrical soundings applied to groundwater investigations, *Geophys. Prospect.,* **41,** 207–227.

Schmutz, M., Albouy, Y., Guerin, R., Maquaire, O., Vassal, J., Schott, J.J. & Descloitres, M., 2000. Joint electrical and time domain electromagnetism (TDEM) data inversion applied to the super sauze earthflow (France), *Surveys Geophys.,* **21,** 371–390.

Vozoff, K. & Jupp, D.L.B., 1975. Joint inversion of geophysical data, *Geophys. J. R. astr. Soc.,* **42,** 977–991.

Williamson, P.R., Sams, M.S. & Worthington, M.H., 1993. Crosshole imaging in anisotropic media, *The Leading Edge,* **12,** 19–23.

Zelt, C.A. & Smith, R.B., 1992. Seismic inversion for 2-D crustal velocity structure, *Geophys. J. Int.,* **108,** 16–34.

Zhang, J. & Morgan, F.D., 1997. Joint seismic and electrical tomography, *Ann. Symp. Environ. Engin. Geophys. Soc. (SAGEEP), Exp. Abst.,* pp. 391–395.

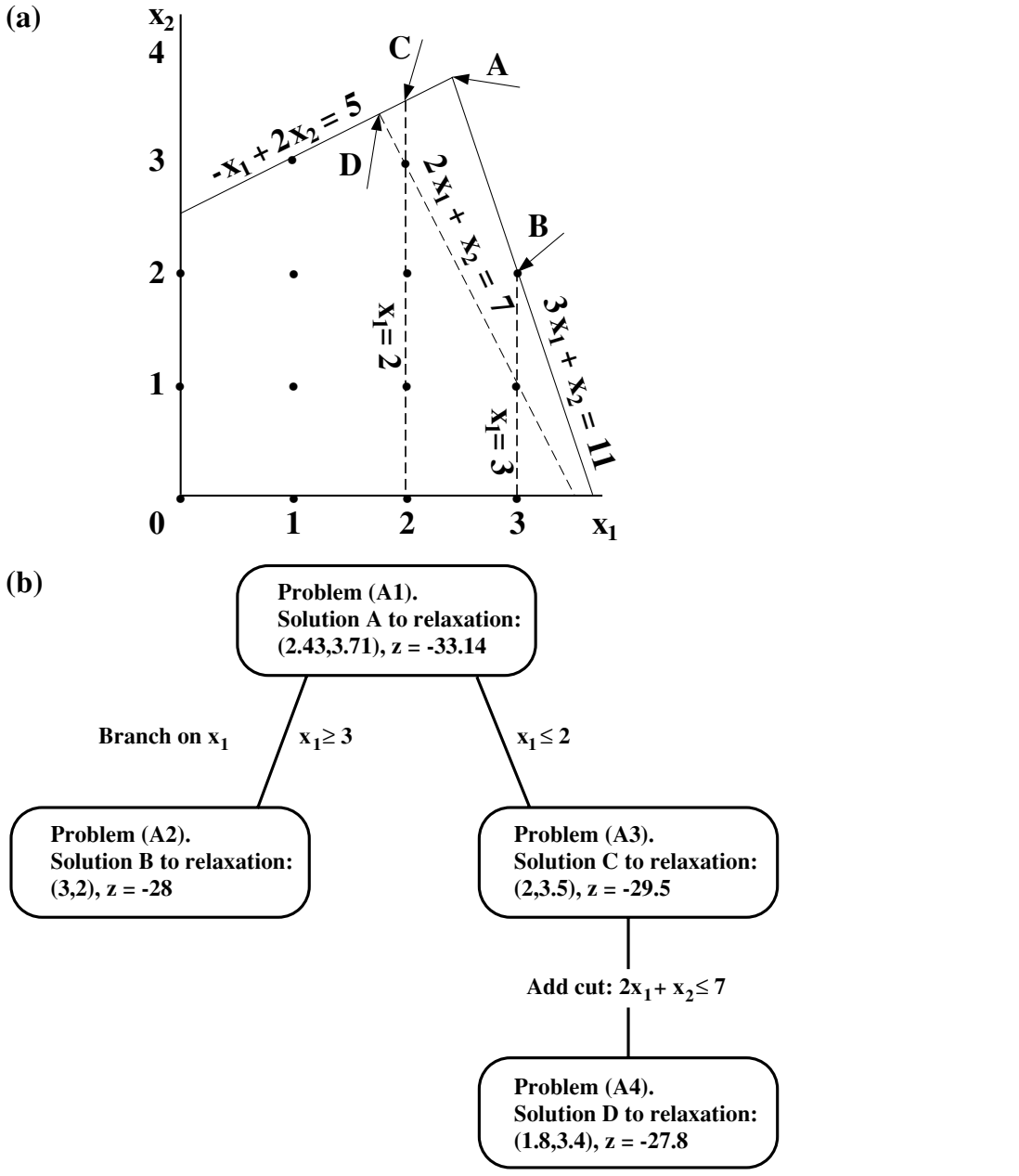

Figure A1. (a) 2-D integer programming problem. The region of continuous solution space is outlined by solid lines. Dots indicate feasible integer solutions. Dashed lines indicate branch-and-cut constraints. (b) Progress of the branch-and-cut process applied to the 2-D integer programming problem by first branching on x_1. Problems A1–A4 and solutions A–D are discussed in the text (modified after Mitchell 2000).

APPENDIX A: MIXED-INTEGER LINEAR PROGRAMMING

Many combinatorial optimization problems can be formulated as mixed-integer linear programming problems. They can be solved by branch-and-cut methods, which are exact algorithms that combine cutting plane with branch-and-bound methods (Land & Doig 1960; Grötschel & Holland 1991; Padberg & Rinaldi 1991; Floudas 1995; Mitchell 2000).

The most widely used technique for solving integer problems is the branch-and-bound method. Subproblems are created by restricting the range of the integer variables. A variable with a lower bound l and an upper bound u is divided into two problems with ranges l to q and $q + 1$ to u, respectively, where q is obtained by rounding the continuous solutions. Lower bounds on the objective function are provided by the linear programming (LP) relaxation to the problem, which involves maintaining the objective function and constraints, but relaxing the integrality restrictions to derive a continuous LP problem. If the optimum solution to a relaxed problem is integral, it is an optimum solution to the subproblem, and the associated value can be used to terminate searches of subproblems that have higher lower bounds.

In the branch-and-cut method, the lower bound of the objective function is again provided by the LP relaxation to the integer problem. The optimum solution to this problem is at a corner of the polyhedron that represents the 'feasible' region. If the optimum

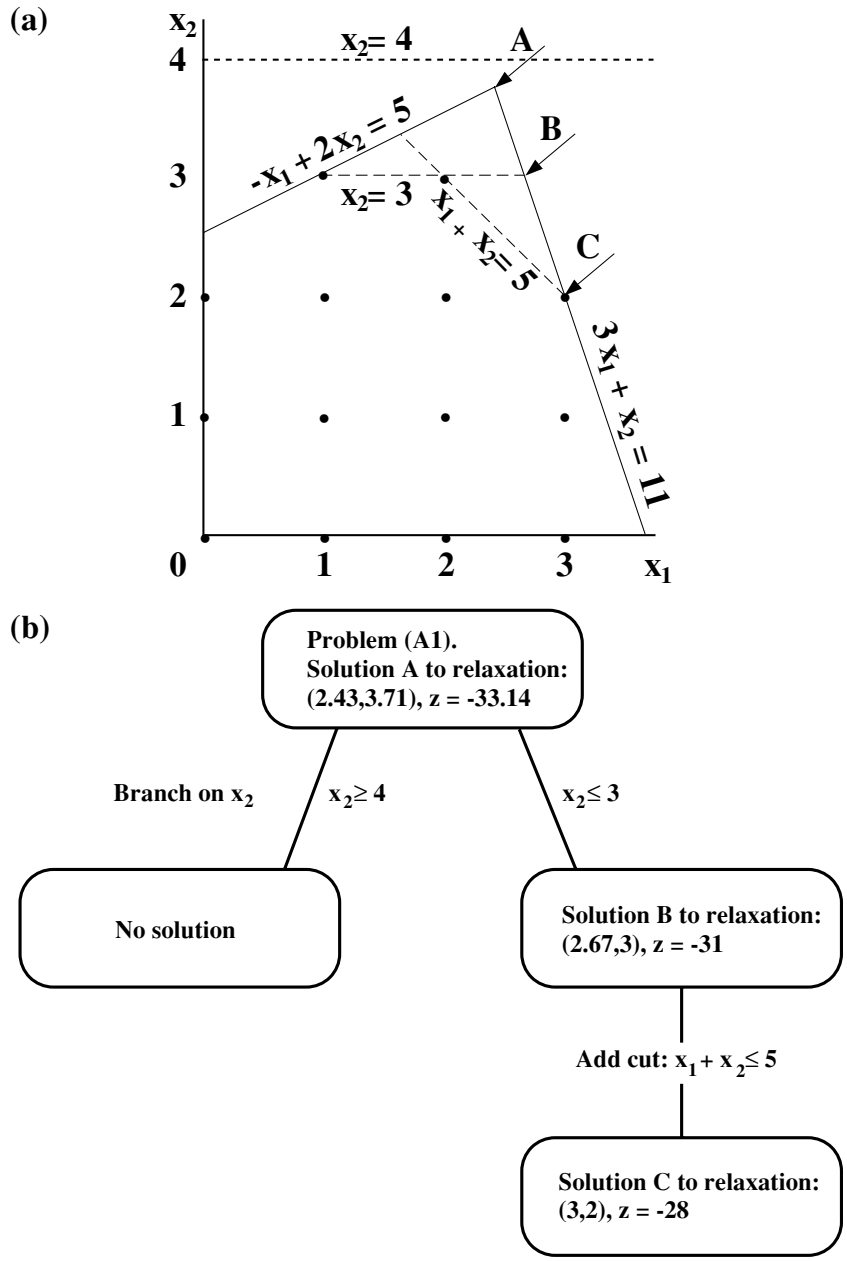

Figure A2. (a) 2-D integer programming problem. The region of continuous solution space is outlined by solid lines. Dots indicate feasible integer solutions. Dashed and dotted lines indicate branch-and-cut constraints. A–C denote possible solutions. (b) Progress of the branch-and-cut process applied to the 2-D integer programming problem by first branching on x_2 (modified after Mitchell 2000).

solution to the LP problem is not integral, this method searches for a constraint that is violated by this solution, but is not violated by any integer solutions. This constraint is called a cutting plane. When this constraint is added to the LP problem, the old optimum solution is no longer valid, so the new optimum solution will be different, potentially providing a better lower bound. Cutting planes are added iteratively until either an integral solution is found or it becomes impossible or too expensive to find another cutting plane. In the latter case, a traditional branch-and-bound operation is performed and the search for cutting planes continues within the subproblems.

To illustrate briefly the basic ideas behind a branch-and-cut algorithm, a simple example with two variables is presented (after Mitchell 2000). The integer programming problem

$$\begin{aligned} \min z := & -6x_1 - 5x_2 \\ \text{subject to} \quad & 3x_1 + x_2 \leq 11 \\ & -x_1 + 2x_2 \leq 5 \\ & x_1, x_2 \geq 0, \text{ integer} \end{aligned} \tag{A1}$$

is illustrated in Fig. A1(a). Possible integer solutions to eq. (A1) are marked with dots. By dropping the integer restrictions, an LP relaxation is obtained. Continuous solutions are contained within the polyhedron outlined by the solid lines.

A branch-and-cut approach first solves the LP relaxation using the simplex algorithm (Press *et al.* 1992), giving the point (2.43, 3.71; A in Fig. A1a) with a value of −33.14. There is now a choice: the LP relaxation can be improved by adding a cutting plane (an inequality that cuts off part of the LP relaxation), or the problem can be divided into two by restricting a variable to be above or below appropriate integer values (i.e. branch-and-bound on x_1 (below and including 2, and above and including 3) or x_2 (below and including 3, and above and including 4)). Note, that these integers are obtained by rounding the solution to the continuous problem.

If the algorithm branches on x_1, two new problems are obtained (Fig. A1b):

$$\begin{aligned} \min z := & -6x_1 - 5x_2 \\ \text{subject to} \quad & 3x_1 + x_2 \leq 11 \\ & -x_1 + 2x_2 \leq 5 \\ & \mathbf{x}_1 \quad \geq 3 \\ & x_1, x_2 \geq 0, \text{ integer.} \end{aligned} \tag{A2}$$

and

$$\begin{aligned} \min z := & -6x_1 - 5x_2 \\ \text{subject to} \quad & 3x_1 + x_2 \leq 11 \\ & -x_1 + 2x_2 \leq 5 \\ & \mathbf{x}_1 \quad \leq 2 \\ & x_1, x_2 \geq 0, \text{ integer.} \end{aligned} \tag{A3}$$

The optimum solution to the original problem will be the better of the solutions to these two subproblems. The solution to the LP relaxation of eq. (A2) is (3, 2; B in Fig. A1a) with a value of −28. This solution is integral, so it solves eq. (A2) and becomes the incumbent best known feasible solution. The LP relaxation of eq. (A3) has an optimum solution of (2, 3.5; C in Fig. A1a) with a value of −29.5. This point is non-integral (it does not solve eq. A3), so that eq. (A3) must be attacked further with additional constraints. For problems with many variables, the strategy would be similar, except the depth of the branching tree may become very large.

Assume a cutting plane that adds the inequality $2x_1 + x_2 \leq 7$ to eq. (A3) (the dashed line in Fig. A1a). This is a valid inequality, in that it is satisfied by every integral point that satisfies eq. (A3). Furthermore, this inequality explicitly excludes (2, 3.5), so it is a cutting plane. The resulting subproblem is

$$\begin{aligned} \min z := & -6x_1 - 5x_2 \\ \text{subject to} \quad & 3x_1 + x_2 \leq 11 \\ & -x_1 + 2x_2 \leq 5 \\ & x_1 \leq 2 \\ & 2\mathbf{x}_1 + \mathbf{x}_2 \leq 7 \\ & x_1, x_2 \geq 0, \text{ integer.} \end{aligned} \tag{A4}$$

The LP relaxation of eq. (A4) has an optimum solution of (1.8, 3.4; D in Fig. A1a) with a value of −27.8. Since the optimum value for this modified relaxation is larger than the value of the incumbent solution, the second subproblem must be at least as large as the value of the relaxation (the optimum solution to the continuous problem). Therefore, the incumbent solution is better than any feasible integral solution for eq. (A4), so it actually solves the original problem. Note, that if the solution of eq. (A4) was lower than that of eq. (A2), another round of branch-and-cut would be performed.

For completeness, the same problem is solved by first branching on x_2 (Fig. A2). For the branch with $x_2 \geq 4$ there is no solution. For the branch with $x_2 \leq 3$, the solution to the LP relaxation is (2.67, 3; B in Fig. A2a) with a value of −31. Introduction of a cutting plane $x_1 + x_2 \leq 5$ results in the solution of (3, 2; C in Fig. A2a) with a value of −28. This solution is integral and since there are no additional subproblems, it solves the integer problem. Note, that the optimum integer solution is found regardless of the sequence of steps. However, one sequence will probably converge with fewer computations than the others.

There are several issues that have to be resolved during the search for an optimum solution. These include the questions of deciding whether to branch or to cut and deciding how to branch and how to generate cutting planes. It is, however, beyond the scope of this article to describe these and related issues for general problems. CPLEX is equipped with various algorithms that are executed throughout the course of the optimization.